

observed discrepancy between two corresponding values are not of a very serious kind. So that in any discussion of the variation of the twenty-three pairs of organs discussed in the present paper, or of the pairs of shrimp organs discussed in my previous communication, it may be assumed as at least an empirical working rule that Galton's function has the same value in all local races. The question whether this empirical rule is rigidly true will have to be determined by fuller investigation, based on larger samples: but the value of a merely empirical expression for the relation between abnormality of one organ and that of another is very great. It cannot be too strongly urged that the problem of animal evolution is essentially a statistical problem: that before we can properly estimate the changes at present going on in a race or species we must know accurately (*a*) the percentage of animals which exhibit a given amount of abnormality with regard to a particular character; (*b*) the degree of abnormality of other organs which accompanies a given abnormality of one; (*c*) the difference between the death rate per cent. in animals of different degrees of abnormality with respect to any organ; (*d*) the abnormality of offspring in terms of the abnormality of parents, and *vice versa*. These are all questions of arithmetic; and when we know the numerical answers to these questions for a number of species we shall know the direction and the rate of change in these species at the present day—a knowledge which is the only legitimate basis for speculations as to their past history and future fate.

### III. "Contributions to the Mathematical Theory of Evolution."

By KARL PEARSON, M.A., Professor of Applied Mathematics,  
University College. Communicated by Professor HENRICI,  
F.R.S. Received October 18, 1893.

(Abstract.)

1. If a series of measurements, physical, biological, anthropological, or economical, not of the same object, but of a group of objects of the same type or family, be made, and a curve be constructed by plotting up the number of times the measurements fall within a given small unit of range to the range, this curve may be termed a *frequency curve*. As a rule this frequency curve takes the well known form of the curve of errors, and such a curve may be termed a *normal frequency curve*. The latter curve is symmetrical about its maximum ordinate. Occasionally, however, frequency curves do not take the normal form, and are then generally, but not necessarily, asymmetrical. Such abnormal curves arise particularly in biological measurements; they have been found by Professor

Weldon, for the measurements of a certain organ in crabs, by Mr. Thompson, for prawns, by Mr. Bateson, for earwigs. They occur, however, in physics, *e.g.*, Dr. Venn's barometric and thermometric frequency curves; in anthropology, *e.g.*, Signor Perozzo's curves for Italian recruits, and Dr. C. Roberts' curves for the eyesight of Marlborough College boys and in fever mortality statistics; in economics, Mr. Edgeworth's curves of prices, and curves I have had drawn for rates of interest.

Frequency curves may, however, be abnormal and yet symmetrical. These are much more likely to deceive even the trained statistician; such curves might arise in target practice, and would be due, for example, to firing with equal precision, *but with a change of sighting* at mid-firing.

2. Abnormal frequency curves fall into three distinct classes :

- a. Asymmetrical curves best represented by a point-binomial, or by its limit a continuous curve.
- b. Asymmetrical curves which are the resultant of two or more normal curves, with different positions of axes, different areas, and different *standard deviations*—a term used in the memoir for what corresponds in frequency curves to the error of mean square.
- c. Symmetrical abnormal curves, which are compounded of two or more normal curves having coincident axes but different areas and standard deviations, or of two normal curves with the same areas and standard deviations but different axes.

3. Let  $\alpha$  be the area of any frequency curve, let the vertical through its centroid, or the line through its centroid perpendicular to the axis of measurement be drawn, and let the second, third, fourth, fifth, and sixth moments about this centroid-vertical,  $\alpha\mu_2$ ,  $\alpha\mu_3$ ,  $\alpha\mu_4$ ,  $\alpha\mu_5$ , and  $\alpha\mu_6$ , be ascertained. This can be done by graphical or arithmetical processes indicated in the memoir, tables being given to assist the calculation in the latter case. Then we can treat the three classes of abnormal curves in the following manner:—

4. *Class a.*—Let the binomial corresponding to the curve be:  $\alpha(p+q)^n$ , where  $p$  = probability in favour of an isolated event,  $q$  = probability against, and  $n$  = number of contributory "causes" in a single trial. For example: the simultaneous spinning of  $n$  teetotums with black and white sides proportional respectively to  $p$  and  $q$ , and  $\alpha$  the total number of times the group of  $n$  is spun. Then it is easy to fit this point-binomial to a frequency curve of which the centroid vertical and  $\mu_1$ ,  $\mu_2$ ,  $\mu_3$ , and  $\mu_4$  are known. The solution for this case is not discussed in the memoir, having been already dealt with by the author.

If it be desired to draw a continuous curve corresponding to the

asymmetrical curve, we can proceed as follows: Imagine  $n$  to be large, but the ratio  $q/p$  either small or large; then we can obtain *generalised form* of the normal curve of an asymmetrical character; its equation referred to the centroid vertical is

$$y = \frac{a}{\sqrt{(2\pi\mu_2)}} \left\{ \frac{\sqrt{(2\pi\beta)} \beta^\beta e^{-\beta}}{\Gamma(\beta+1)} \right\} \left( 1 + \frac{\mu_3}{2\mu_2^2} x \right)^{\beta-1} e^{-\frac{2\mu_2}{\mu_3} x},$$

where  $\beta$  stands for  $4\mu_3^3/\mu_2^3$  and  $\Gamma(\beta)$  is the Eulerian gamma function.\* Putting  $\mu_3 = 0$  for an asymmetrical curve, the equation takes an indeterminate form obtained by putting  $\beta = \infty$ , but on evaluation we have the usual normal form :

$$y = \frac{a}{\sqrt{(2\pi\mu_2)}} e^{-\frac{x^2}{2\mu_2}}.$$

This generalised probability curve fits with a high degree of accuracy a number of measurements and observations hitherto not reduced to theoretical treatment, *e.g.*, barometric frequency curves.

The importance of this first dissection of asymmetrical frequency curves lies in the fact that it measures the theoretical number  $n$  of contributory "causes" and the odds  $p:q$  that an element of deviation will be *positive*. The whole theory is, however, of an elementary character, and, as biological frequency curves often tend to develop a double-humped character,† they do not invariably fall under this class, and it is not dealt with at length in the memoir.

5. *Class b.*—The general theory of the dissection of a given abnormal frequency curve into  $m$  components is not dealt with, partly on account of its exceedingly great analytical difficulties, partly because there is an *a priori* probability that we have a mixture of only two homogeneous groups, or from the standpoint of evolution that the species will break up at first into two, rather than three or more, families. At any rate, the dissection into two is likely to give us either the chief components or a measure of the chief asymmetry of the curve. Supposing the curve asymmetrical, it is shown that the solution of the problem is theoretically unique, but it is pointed out that in statistical practice our curve is based upon a limited number of measurements, and is therefore not an accurately true compound of two normal groups. A theoretical test is given to distinguish between the better of two or more solutions. The method adopted for the dissection is based on equality of the first five moments and of the areas of the abnormal curve and of its two components. This method is justified in the same manner as the determination of the

\* If  $\beta$  be large, it may be taken as approximately whole, and the factor in round brackets is then unity.

† *E.g.*, claspers of earwigs, height of Italian recruits of various special provinces, short sight of Marlborough boys, height of inhabitants of Doubs, &c.

normal curve by fitting any series of observations by aid of the area and the first two moments (*i.e.*, the first moment gives the mean, and the second the error of mean square) is justified. The method leads to what is termed the *fundamental nonic*, every root of which gives a real or imaginary solution of the problem. The best solution is selected by the criterion that it gives the closest approach to the given frequency curve in the value of the sixth moment. From the nonic is deduced a quadratic for the areas of the components corresponding to each solution. If both roots of this quadratic are real and positive, we have either a mixture of two heterogeneous species, or evolution is breaking the homogeneous material up into two families of different magnitudes, different means, and different standard deviations from the mean. If one root of the quadratic be real and positive, and the other real and negative, we have evolution destroying a certain percentage round a certain mean out of an initially homogeneous and normal group.

Should one of the standard deviations be imaginary, we get the percentage of anomalous and irregular measurements in a homogeneous group.

6. *Class c.*—The solution here is unique and depends upon the equality of the areas and of the first *six* moments; for all odd moments vanish, and we have four quantities to determine, *i.e.*, the percentages of each group and their standard deviations. The solution depends on a quadratic for the areas, and the same remarks apply as to the quadratic for Class *b*.

7. Rules are given for detecting whether we have a mixture of two groups, or whether a differentiation into species of a homogeneous material is going on; and also rules for measuring the amount of asymmetry which is to be considered significant. The former rules are, briefly:—

- i. Select the most asymmetrical curve out of the curves for the organs measured; dissect it into two curves or groups by the method for Class *b*.
- ii. Select the most symmetrical curve out of the curves for the organs measured and dissect it into two groups by the method for Class *c*, or, if it have significant asymmetry by the method for Class *b* again. Then ( $\alpha$ ), if the first dissection is possible and the second is not, a real evolution is going on; ( $\beta$ ), if the first dissection is possible and the second is possible, and both groups give sensibly the same percentages, we have a mixture of two heterogeneous materials and no true evolution, unless the organs be so closely allied that one must vary directly with the other (*e.g.*, length of right and left legs); ( $\gamma$ ), both dissections are possible, but give groups with different percentages; we have *both* organs evolving differently at the same time.

8. The theory is applied to Professor Weldon's measurements on Naples crabs. It is shown that his material is absolutely homogeneous, all roots of the nonic for No. 4 organ leading to imaginary solutions, even its real root. On the other hand, it is shown that the Naples crabs are breaking up into two different sized families, owing to evolution in their foreheads. The theory is further applied to Mr. Thompson's measurements of the carapace of prawns (1,000 measurements). It is shown that we have in the measurements a very small percentage of anomalous results, corresponding to prawns deformed in this organ, or that there is, on the other hand, a small but unstable giant population mixed with the normal population. Which of these results is to be considered the true answer to the problem can only be determined after an analysis of the frequency curves for other organs.

From the mathematical standpoint, the memoir illustrates the determination of the roots of equations of the ninth degree, and the calculation of the higher moments of curves.

#### IV. "Experiments in Heliotropism." By G. J. ROMANES, F.R.S. Received October 2, 1893.

I cannot find in the literature of heliotropism that any experiments have hitherto been made on the effects of interrupted illumination, when the periods of illumination are rendered as brief as possible—*i.e.*, instantaneous flashes of light. Accordingly I have conducted an extensive research on heliotropism, where the flashes have been caused either by means of electric sparks in a dark room, or by the opening of a photographic shutter placed before the plants in a camera obscura with an arc light or Swan burner, at a distance of several feet on the other side of the shutter. The electric sparks were made either with a Wimshurst machine, induction sparks, or by means of the following contrivance. From the binding screws of the condenser of a large induction coil copper wires were led to a cup of mercury, where, by means of an electro-magnet suitably actuated by clock work, a current was closed and opened at any desired intervals: each break was therefore accompanied by a brilliant spark. A thick plate of glass was interposed between the seedlings and the electrical apparatus. In all the experiments here described the plants employed were mustard seedlings (*Sinapis nigra*), previously grown in the dark until they had reached a height of between 1 and 2 inches. Save when the contrary is stated, in all the experiments comparative estimates were formed by using the same pot of seedlings: during the first half of a comparative experiment half of the seedlings were protected from the light by a cap of cardboard covering half the pot;